Existence of (cheap) minimal norm controllers

Engwerda, J.C.

Published in:
International Journal of Control

Publication date:
1989

INT. J. CONTROL, 1989, VOL. 50, NO. 3, 755–762

Existence of (cheap) minimal norm controllers

J. C. ENGWERDA†

It is shown that the spectral norm of the closed-loop system matrix is minimized if a special type of minimum variance control is applied. Furthermore, a sufficient condition is derived for the existence of a controller which, with less control effort than this minimum variance controller, obtains the same minimal closed-loop norm.

1. Introduction

In both economics and industry there exist many systems for which the evolution of the system parameters in the future is unknown and, moreover, information about the state of the system is not available in time. Due to this lack of information the problem of stabilizing such systems is a difficult one. However, if the current system parameters are known to the designer, he can use this information to design a stabilizing linear state-feedback controller. This approach will be successful if the operator norm of the resulting closed-loop (CL) system matrix is uniformly smaller than one at any current time instant. For this reason we concentrate in this work on finding linear state-feedback controllers which, at any time instant, minimize the operator norm of the current CL system matrix.

Now, in general it is not possible to design a state feedback controller which yields a minimal norm of the CL system that is smaller than one. This is a direct consequence of the Eckart–Young theorem (see § 2). So, in these situations it is doubtful whether the system will be stabilized if one sticks to this policy of controlling the system. However, what is the alternative if really no information is available concerning the future development of the system parameters? Of course, one can guess at the parameter developments, but from this point of view, the judgment that the system parameters, e.g., will be time-invariant is as good (or bad) as the judgment that the system has some predescribed time-varying behaviour. Therefore the policy of minimizing the norm of the current CL system matrix, if the future parameter developments are unknown is, given the circumstances, the most rational one. This is in spite of the fact that, when performing this policy, it may occur that we obtain an unstable CL system though the system is stabilizable (as it turns out after all).

Now, the problem of finding a state feedback controller yielding a minimal CL spectral norm has been studied in operator theory in a more general setting. This is usually solved by translating it to a dilation problem (see, e.g., Doyle 1984, p. 56). However, in general there exists a whole class of controllers which all yield this minimal norm. So a natural question which thus arises is whether it is possible to select among this class of controllers a controller which attains this minimal norm with as little control effort as possible. This is the main object of study in this work.

To tackle this problem we first treat in § 2 the original problem of finding a controller which gives rise to a minimal norm, from a geometric point of view. It turns

Received 12 July 1988. Revised 18 October 1988.

† Department of Econometrics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

out that a special type of minimum variance control always yields this minimal obtainable norm. Since this minimum variance controller also minimizes a quadratic cost criterion in which the control costs play no role, this enables an intuitive approach to the main problem. For, the objective to be obtained with less control effort, the same minimal norm of the CL system matrix can now be modelled by introducing in the cost criterion an additional term penalizing quadratically controlled efforts.

The amount of control penalty in this extended cost criterion depends on a positive definite weighting matrix R. Consequently, the controller minimizing this new cost criterion depends on R too. In § 3 we state necessary and sufficient conditions for the existence of a less expensive control policy among this class of controllers parametrized by R. The work ends with a conclusion.

The results we state here have partly been reported by Engwerda (1987 a).

2. Minimal norm controller

In this section, we derive a linear state feedback controller which yields a minimal spectral norm of the closed-loop matrix at any point in time. The motivation to study just the spectral norm is due to the fact that for any matrix norm we have that $\|A\| > r(A)$. Here $r(A)$ denotes the spectral radius of matrix A (i.e. the absolute largest eigenvalue of A). Moreover, we have for normal matrices (matrices for which $AA^{\mathrm{T}} = A^{\mathrm{T}}A$) that the spectral norm equals the spectral radius (see Lancaster and Tismenetsky 1985, p. 359). So for normal matrices, the spectral norm measures the smallest norm that can be obtained for a matrix. Note that this implies for time-invariant systems that the asymptotic stability property of the CL-system is improved as much as possible if the state feedback minimizes the spectral norm of the CL-system matrix.

Before we turn to the main point of this section, we introduce first the system and some notation. The system we consider in this work is described by the following first-order difference equation:

$$\Sigma: \quad x(k+1) = A(k)x(k) + B(k)u(k), \quad x(0) = x$$

Here $x(k) \in \mathbb{R}^n$ is the state of the system and $u(k) \in \mathbb{R}^m$ is the applied control. Let $\sigma'(A) := \sigma_1(A) \geqslant \ldots \geqslant \sigma_n(A)$ denote the singular values of matrix A. Then, the spectral norm of A is defined as the largest singular value of this matrix, i.e. $\|A\|_s = \sigma_1(A)$. By $\|A\|_E := \left(\sum_{i,j} a_{ij}^2\right)^{1/2}$ we denote the euclidean norm of matrix A. Finally, A^{T} denotes the transpose of matrix A, and A^+ its Moore–Penrose inverse (see Lancaster and Tismenetsky 1985, § 12.8).

The following proposition is well known and can be found, e.g., in work by Basilevsky (1983, p. 240).

Proposition 1

$$\|A\|_s = \max_{x \neq 0} \frac{\|Ax\|_E}{\|x\|_E}$$

This property is also often used to define the spectral norm of a matrix. The next property tells us how well a matrix can be approximated in norm by another matrix with a predescribed rank. It is known as the Eckart–Young theorem.

Proposition 2

$$\min_F \{\|A - F\|_s \mid \operatorname{rank}(F) = p\} = \sigma_{p+1}(A)$$

Consequently the spectral norm of the closed-loop matrix $\|A + BF\|_s$ is for any matrix F, at least $\sigma_{m+1}(A)$. So, in general, it will not be possible to design for a system Σ a feedback controller which has the property that the norm of the closed-loop matrix is smaller than one.

The next lemma is a preparation for the main result of this section. A proof of it can be found in work by Lancaster and Tismenetsky (1985, § 12.9).

Lemma 1

Let $z \in \mathbb{R}^n$ and $B \in \mathbb{R}^{n \times m}$ be given. Then, $\min_y \|z + By\|_E$ is obtained by $y = -B^+z$.

The main result reads as follows.

Theorem 1

$\min_F \|A + BF\|_s$ is obtained by $F = -B^+A$.

Proof

According to Proposition 1 and Lemma 1 we have that

$$\|A + BF\|_s = \max_{x \neq 0} \frac{\|(A + BF)x\|_E}{\|x\|_E} \geqslant \max_{x \neq 0} \frac{\|(I - BB^+)Ax\|_E}{\|x\|_E}$$

$$= \|(I - BB^+)A\|_s$$

So, $\min \|A + BF\|_s \geqslant \|(I - BB^+)A\|_s$.

From this last inequality the statement of the theorem is immediate. □

In the remainder of this work we will denote the matrix $I - BB^+$ by P.

3. Existence conditions for a cheaper controller

In Theorem 1 we derived a controller which always gives rise to a minimal spectral norm of the closed-loop system matrix. The feedback gain was given by $-B^+A$. Now, this controller is also obtained if we take $S = I$ in the following minimum variance optimization problem.

Proposition 5

Consider $J(k) = x^{\mathrm{T}}(k)Qx(k)$, with $Q = S^{\mathrm{T}}S$ for some square matrix S. Then,

$$\min_{u(k)} J(k+1) \quad \text{subject to } \Sigma$$

is obtained by $u_0(k) = -(SB)^+SAx(k)$.

So, if we are looking for a controller which yields the same minimal norm with less control effort, it seems reasonable to consider the following cost functional:

$$J'(k+1) = x^{\mathrm{T}}(k+1)x(k+1) + u^{\mathrm{T}}(k)R(k)u(k) \quad \text{with } R(\cdot) > 0$$

Here the notation $M > (=) 0$ means that matrix M is (semi) positive definite.

The idea is that the effect of minimizing J' instead of J w.r.t. Σ will probably be that the amount of applied control is less the more positive matrix R is, provided $B^{\mathrm{T}}A$ differs from zero. That this conjecture is only partly true is shown in the next proposition and example. However, first we give a proper definition of what is meant by a cheaper controller.

Definition 1

A controller $u_1 = F_1 x$ is called cheaper than $u_2 = F_2 x$ if $\|A + BF_1\|_s = \|A + BF_2\|_s = \|PA\|_s$, whereas $\|u_1\|_E \leqslant \|u_2\|_E$ for any initial state x of Σ and strict inequality holds for at least one initial state.

Proposition 4

$\min_u J'(k+1)$ w.r.t. Σ is obtained by $u'(k) = -(R + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Ax(k)$.

Moreover we have that for any initial state x of the system at time k, $u'(k) = u_0(k)$ iff $B^{\mathrm{T}}A = 0$.

Proof

The proof of the first part of the statement can be found, e.g., in work by Kirk (1970). To prove the second part of the proposition we only show that from the consideration that $u'(k)$ equals $u_0(k)$, for any initial state, it follows that $B^{\mathrm{T}}A$ equals zero. The reverse implication is trivial.

To this end we first note that $B^{\mathrm{T}} = B^{\mathrm{T}}BB^{+}$. Consequently $B^{+}A = (R + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}A$ iff

$$B^{+}A = (R + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}BB^{+}A$$

Rewriting this equality yields

$$-(R + B^{\mathrm{T}}B)^{-1}RB^{+}A = 0$$

Since, by assumption, R is positive definite it follows that $B^{+}A$ must be zero.

This completes the argument. □

In the sequel, we shall abbreviate the closed-loop matrix, which results if the optimal control $u'(k)$ is applied, by $P'(k)A(k)$.

The next example originates from the fact that $0 < R_1 \leqslant R_2$ does not generally imply that $R_1^2 \leqslant R_2^2$.

Example 1

Let

$$A = I, \quad B = R_1^{1/2} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\frac{1}{8}} \end{bmatrix}, \quad R_2 = \begin{bmatrix} 3 & -2 \\ -2 & 2\frac{3}{8} \end{bmatrix}$$

Then $0 < R_1 \leqslant R_2$, but $\|(R_1 + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Ax\|_E \not\geqslant \|(R_2 + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Ax\|_E$ for all x.

Despite this negative result, we are able to derive a condition for the existence of a cheaper controller of the form $u'(k)$. This existence question is basically answered in Theorem 2, but before we make a start in solving this problem we first provide two examples. One example shows that it is sometimes possible to find a cheaper control than $-B^{\mathrm{T}}Ax$; the other deals with a situation in which this is not possible.

Example 2

Let

$$A = \begin{bmatrix} 4 & 2 \\ -1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad R = 1$$

Then $\|PA\|_s = \|P'A\|_s = \sqrt{5}$.

Example 3

Let B be invertible and $A \neq 0$. Then $\|PA\|_s = 0$ and $\|P'A\|_s \neq 0$ for any $R > 0$.

The following two lemmas are a preparation for the main theorem of this section.

Lemma 2

Let

$$D := \begin{bmatrix} E & G^{\mathrm{T}} \\ G & H \end{bmatrix} \geqslant 0$$

with E and H square matrices. Moreover, assume that $\max(\|E\|_s^2, \|H\|_s^2) = \|H\|_s^2 = \sigma^2$. Let $\mathbf{H}$ be the eigenspace of H corresponding to σ. Then, $\exists\, x \in \mathbf{H}$ such that $x^{\mathrm{T}}G \neq 0 \Rightarrow \|D\|_s^2 > \sigma^2$.

Proof

If $\sigma = 0$, the statement is trivial. Therefore assume $\sigma \neq 0$. Let $\bar{y}$ be an eigenvector from H such that $\bar{y}^{\mathrm{T}}G \neq 0$. Let $\bar{x}^{\mathrm{T}} = (1/\sigma)\bar{y}^{\mathrm{T}}G$. Then

$$\|D\|_s = \max_{(x,y)\neq 0} \frac{x^{\mathrm{T}}Ex + 2y^{\mathrm{T}}Gx + y^{\mathrm{T}}Hy}{\|x\|_E^2 + \|y\|_E^2} \geqslant \frac{\bar{x}^{\mathrm{T}}E\bar{x} + 2\bar{y}^{\mathrm{T}}G\bar{x} + \bar{y}^{\mathrm{T}}H\bar{y}}{\|\bar{x}\|_E^2 + \|\bar{y}\|_E^2}$$

$$= \frac{\dfrac{1}{\sigma^2}\bar{y}^{\mathrm{T}}GEG^{\mathrm{T}}\bar{y} + \dfrac{2}{\sigma}\bar{y}^{\mathrm{T}}GG^{\mathrm{T}}\bar{y} + \sigma\bar{y}^{\mathrm{T}}\bar{y}}{\dfrac{1}{\sigma^2}\bar{y}^{\mathrm{T}}GG^{\mathrm{T}}\bar{y} + \bar{y}^{\mathrm{T}}\bar{y}}$$

$$= \sigma + \frac{1}{\sigma}\bar{y}^{\mathrm{T}}GG^{\mathrm{T}}\bar{y} + \frac{\bar{y}^{\mathrm{T}}GEG^{\mathrm{T}}\bar{y}}{\bar{y}^{\mathrm{T}}GG^{\mathrm{T}}\bar{y} + \sigma^2\bar{y}^{\mathrm{T}}\bar{y}}$$

Since $\bar{y}^{\mathrm{T}}G \neq 0$ and $E \geqslant 0$, it is clear from this last equation that $\|D\|_s$ is strictly greater than σ which proves the lemma. □

Lemma 3

Let

$$D := \begin{bmatrix} E & G^{\mathrm{T}} \\ G & H \end{bmatrix} \geqslant 0$$

and λ be an eigenvalue of H with multiplicity p. Assume that $\mathbf{H}$ is the eigenspace of H corresponding to λ. Then, if $\forall x \in \mathbf{H}$, $x^{\mathrm{T}} G = 0$, λ is also an eigenvalue of D with multiplicity greater than or equal to p.

Proof

The proof follows immediately from the demonstration that if x is an eigenvector of H that goes with λ, then $\begin{bmatrix} 0 \\ x \end{bmatrix}$ is an eigenvector of D that goes with λ. □

For the formulation of the theorem it is convenient to choose another orthonormal basis in $\mathbb{R}^n$. From now on we assume that $\mathbb{R}^n$ is decomposed as $\operatorname{Im} B(k) \oplus X(k)$. With respect to this basis, matrix $B(k)$ equals $\begin{bmatrix} B' & 0 \\ 0 & 0 \end{bmatrix}$, where B' is a square $m' \times m'$ invertible matrix. Furthermore, we write

$$A(k) := \begin{bmatrix} E(k) & F(k) \\ G(k) & H(k) \end{bmatrix}$$

where $E(k) \in \mathbb{R}^{m' \times m'}$ and $H(k) \in \mathbb{R}^{(n-m) \times (n-m')}$. Note that due to this choice of the basis, the cost criterion does not change. The theorem then becomes as follows.

Theorem 2

Let $\|PA\|_s = \sigma$, and $\mathbf{H}$ be the eigenspace of $PAA^{\mathrm{T}}P$ belonging to σ. Then, the following hold:

(a) if $\sigma \neq 0$, $\exists R > 0$ such that $\|P'A\|_s = \sigma \Leftrightarrow \mathbf{H}(GE^{\mathrm{T}} + HF^{\mathrm{T}}) = 0$; and

(b) if $\sigma = 0$, $\exists R > 0$ such that $\|P'A\|_s = 0 \Leftrightarrow A = 0$.

Proof

Sufficiency. Straightforward calculation shows that with $R^{-1} := \begin{bmatrix} R_{11}^{-1} & R_{12} \\ R_{12}^{\mathrm{T}} & R_{22}^{-1} \end{bmatrix}$,

$$\|P'A\|_s^2 = \|(I - B(R + B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}})A\|_s = \|(I + BR^{-1}B^{\mathrm{T}})^{-1}A\|_s$$

$$= \left\| \begin{bmatrix} (I + B'R_{11}^{-1}B'^{\mathrm{T}})^{-1}E & (I + B'R_{11}^{-1}B'^{\mathrm{T}})^{-1}F \\ G & H \end{bmatrix} \right\|_s$$

$$= r \begin{bmatrix} (I + B'R_{11}^{-1}B'^{\mathrm{T}})^{-1}(EE^{\mathrm{T}} + FF^{\mathrm{T}})(I + B'R_{11}^{-1}B'^{\mathrm{T}})^{-1} & (I + B'R_{11}^{-1}B'^{\mathrm{T}})^{-1}(EH^{\mathrm{T}} + FH^{\mathrm{T}}) \\ (GE^{\mathrm{T}} + HF^{\mathrm{T}})(I + B'R_{11}^{-1}B'^{\mathrm{T}}) & GG^{\mathrm{T}} + HH^{\mathrm{T}} \end{bmatrix}$$

and

$$\|PA\|_s^2 = r \begin{bmatrix} 0 & 0 \\ 0 & GG^{\mathrm{T}} + HH^{\mathrm{T}} \end{bmatrix}$$

Applying Lemma 2 yields that $\mathbf{H}(GE^{\mathrm{T}}+HF^{\mathrm{T}})=0$ is a necessary condition that must be satisfied if $\|PA\|_s=\|P'A\|_s$. Moreover, if $\sigma=0$ necessarily $EE^{\mathrm{T}}+FF^{\mathrm{T}}$ and $GG^{\mathrm{T}}+HH^{\mathrm{T}}$ must be zero. Consequently in this case $A=0$.

Necessity. First we consider the case $\sigma\neq 0$. The spectrum of PA equals $\{0\}\cup\text{spectrum}\,(GG^{\mathrm{T}}+HH^{\mathrm{T}})$. Since the spectrum of a matrix depends continuously on its parameters, the spectrum of the following matrix will be about the same provided S, T, and U are small enough: $\begin{bmatrix} S & T \\ U & GG^{\mathrm{T}}+HH^{\mathrm{T}} \end{bmatrix}$.

Now choose $R=\varepsilon\begin{bmatrix} B'^{\mathrm{T}}B' & 0 \\ 0 & I \end{bmatrix}$, with ε small enough, in $P'AA^{\mathrm{T}}P'$. Applying Lemma 3 we see that the spectrum of this matrix consists of $\{\sigma^2\}$, small perturbations of $\{0\}$ and small perturbations of the other (smaller) eigenvalues of $P'AA^{\mathrm{T}}P'$. Therefore $r(P'AA^{\mathrm{T}}P')=\sigma^2$, which had to be proved.

Finally, we consider the case $\sigma=0$, i.e. $A=0$. In this case both $EE^{\mathrm{T}}+FF^{\mathrm{T}}$ and $GG^{\mathrm{T}}+HH^{\mathrm{T}}$ are zero too. So, it is clear that now any $R>0$ gives rise to a closed-loop matrix with zero norm. □

From the construction part of the theorem we notice that if the existence condition is satisfied in the non-trivial case, i.e. $\sigma\neq 0$, then there always exists a cheaper control provided $B^{\mathrm{T}}A\neq 0$. For

$$B(R+B^{\mathrm{T}}B)^{-1}(R+B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}=\begin{bmatrix} (1+\varepsilon)^{-2}B'^{-\mathrm{T}}B'^{-1} & 0 \\ 0 & 0 \end{bmatrix}\leqslant\begin{bmatrix} B'^{-\mathrm{T}}B'^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$
$$=B^{+\mathrm{T}}B^{+}$$

We shall formulate this result in a corollary.

Corollary 1

With the notation given in Theorem 2, there exists a cheaper control of the type $u'=-(R+B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Ax$, where $R>0$, iff the following two conditions are satisfied:

(*a*) $$B^{\mathrm{T}}A\neq 0$$

(*b*) $$\mathbf{H}(GE^{\mathrm{T}}+HF^{\mathrm{T}})=0$$

Note that this corollary only gives an existence result.

An interesting question which remains to be solved is, of course, how in practice the weighting matrix R can be determined, if it exists, that yields the cheapest control among this class of controllers. That indeed a cheapest control sometimes exists can be seen from Example 2. Easy calculations show that it is obtained by taking $R=3$ (see Example 2).

We conclude this section by remarking that the existence condition is related to a special type of controller. In other words, there may still exist other controllers which are even cheaper than any sample belonging to the class of controllers which satisfy the conditions of Corollary 1. So, Corollary 1 states, in general, merely sufficient conditions for the existence of a cheaper controller. Only in one case it is immediately clear what the overall cheapest controller is. This is when $\|PA\|_s=\|A\|_s$. Then $u=0$ is the cheapest control.

Conclusions

In this work, we have been concerned with minimal norm controllers. It is well known that these controllers exist and that they are, in general, not uniquely determined. Therefore, we have introduced a partial ordering among this class of controllers and have addressed the question of finding a minimal one. This turns out to be a difficult problem. The more so, since the partial ordering introduced does not induce a unique minimal element (in general).

For these reasons, we have restricted the class of controllers among which we were looking for a minimal element. The class of controllers we have considered are the extended minimum variance controllers, since we have seen that a special type of minimum variance control always yields a minimal norm controller. It turns out that among the class of extended minimum variance controllers there sometimes exist controllers which with less control effort achieve the minimal norm. We have given necessary and sufficient conditions for the existence of such a cheaper control. It may be clear that this result is only a first step in solving the general problem.

REFERENCES

BASILEVSKY, A., 1983, *Applied Matrix Algebra in the Statistical Sciences* (Amsterdam: North Holland).

DOYLE, J., 1984, *Advances in Multivariable Control*, Lecture Notes, ONR/Honeywell workshop, Minneapolis.

ENGWERDA, J. C., 1987 a, Existence and robustness of minimal norm controllers. *Proceedings of the International Conference 'Control 88'*, organized by I.E.E., Oxford, pp. 582–586.

KIRK, D. E., 1970, *Optimal Control Theory* (Englewood Cliffs, NJ: Prentice Hall).

LANCASTER, P., and TISMENETSKY, M., 1985, *The Theory of Matrices*, second edition (London: Academic Press).